



## Isolation and characterization of the mouse cystatin B gene.

L A Pennacchio and R M Myers

*Genome Res.* 1996 6: 1103-1109

Access the most recent version at doi:[10.1101/gr.6.11.1103](https://doi.org/10.1101/gr.6.11.1103)

---

**References** This article cites 25 articles, 3 of which can be accessed free at:  
<http://genome.cshlp.org/content/6/11/1103.full.html#ref-list-1>

### License

**Email Alerting Service** Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).



---

To subscribe to *Genome Research* go to:  
<https://genome.cshlp.org/subscriptions>

---

Copyright © Cold Spring Harbor Laboratory Press

## LETTER

# Isolation and Characterization of the Mouse Cystatin B Gene

Len A. Pennacchio<sup>1,2</sup> and Richard M. Myers<sup>2,3</sup>

<sup>1</sup>Department of Biological Sciences and <sup>2</sup>Department of Genetics, Stanford University School of Medicine, Stanford, California 94305-5120

The cystatins make up a large superfamily of proteins that inhibit cysteine proteases. Recently, we showed that loss-of-function mutations in the human cystatin B gene are responsible for progressive myoclonus epilepsy of the Unverricht–Lundborg type (EPM1). However, despite the known role of cystatin B in cysteine protease inhibition, it is not clear why decreased levels of this protein cause EPM1. To provide new insights into the biochemical and pathological mechanisms of EPM1, we are working toward developing an animal model for this disease. Here we present the mouse cystatin B nucleotide and amino acid sequence. We show that the mouse gene spans a 3-kb genomic region and contains 3 exons and 2 introns, identical to the structure of both the rat and human cystatin B genes. The amino acid sequence identity of the protein is 86%, 79%, and 71% to that of the rat, human, and bovine cystatin B proteins, respectively. In addition, we show that the mouse cystatin B gene is expressed in many tissues, similar to results observed previously in humans. Finally, we report the mapping of the mouse cystatin B gene (*Stfb*) to chromosome 10, further extending the synteny between this region of the mouse chromosome and human chromosome 21q22.3.

[The sequence data described in this paper have been submitted to GenBank under accession no. U59807, and the mapping data is available at Mouse Genome Database accession no. MGD-CREX-660.]

The cystatins are a large class of proteins that function as inhibitors of cysteine proteases, including the cathepsins. Structural and functional comparisons suggest that the cystatins comprise a single evolutionary superfamily, presumably derived from a common ancestor (Rawlings and Barrett 1990). This superfamily is classified into three distinct families of closely related proteins—the intracellular stefins (family-1), the secretory cystatins (family-2), and the kininogens (family-3) (Barrett et al. 1986a). Family-1 cystatins are ~100-amino-acid single-chain proteins that lack disulfide bonds.

Of the family-1 type cystatins, stefin A and cystatin B (also known as stefin B) have been studied the most extensively. Both are 98-amino-acid proteins that have been identified in several different mammalian species. Stefin A is found in high concentrations in various types of epithelial cells and polymorphonuclear leukocytes in the liver (Barrett et al. 1986b). In contrast, cystatin B is widely distributed among different cell types

and tissues (Barrett et al. 1986b; Pennacchio et al. 1996). Although stefins lack an export signal sequence and are generally thought to function intracellularly, they have also been found in extracellular fluid (Abrahamson et al. 1986). The amino acid sequences of human, rat, and bovine cystatin B are known (Ritonja et al. 1985; Sato et al. 1990; Turk et al. 1992), and the X-ray structure of the human protein complexed with carboxymethyl-papain has been determined (Stubbs et al. 1990b).

Recently, we showed that defects in the human cystatin B gene are responsible for progressive myoclonus epilepsy of the Unverricht–Lundborg type (EPM1). EPM1 is inherited as an autosomal recessive disease characterized by severe stimulus-sensitive myoclonus, tonic-clonic seizures, and progressive neurological deterioration (Berkovic et al. 1985). The cystatin B gene is expressed ubiquitously and its gene product is thought to have a general role in the inactivation of proteases that escape from the lysosome (Turk and Bode 1991). It is not clear how decreased levels of the ubiquitous cystatin B protein cause the symptoms of EPM1, which are apparently

<sup>3</sup>Corresponding author.  
E-MAIL [myers@shgc.stanford.edu](mailto:myers@shgc.stanford.edu); FAX (415) 725-9689.

## PENNACCHIO AND MYERS

```

ccccacttaatagtcacaggacaaggggaaattgagatgtccagggtgtacttctgtc
-420 cacctactccatgaaagactaggacccttctcacatctctgcatcctgtgaggaccg
gaagcccctaaagccgggtcagagatactctacaatcacttcaacgaccattgaggggcaa
-300 tcaggactcccagcttcaaaacccttcttagcgtccgttcagggaaggacgtgaagcac
ccaccatgcctcaccctccgcccacgcccagcctgggtctcctgactctgcctggaacgg
-180 aggggtggtgacagtaaaagagagcagcagagagggtgcccggcccggccctgctctgt
cacgtgctacgggtccggggggggagcgaagccagggttttctagggtccaggcacc
-60 aggtctcctagttggatctgtcttcagcttctcctgctacccccactactgctgccaag

1 ATGATGTGTGGCGGCCCATCTGCCACAATGCCAGCCACGGCCGAGACGCAGGAGGTTCGCC
M M C G A P S A T M P A T A E T Q E V A

61 GACCAGgtgaggctgggcccagggtcaggccagtctgagccaggcctgcccagaccggcg
D Q

121 gcctcagggaacggcctgccagactgggtcaggctcctgagcgggtttccggggcgccca
caagtgtgactgggagctgggggctctggatctgggtggaagattcagggccgggaactg
241 gggcagcttctcctccgcttgcatcaaaagggccactcacctattaggaaactgcccc
gggatcgggtggaggatccgggtgggtcccagagaattcaggaaggcagtgtagaacctag
361 acggcacccttttctgacttaccccagggcctaaacaagagaagccagactgggtactgtg
gctgtctcctcaaaagaaagactaggactgttagctcagtgccagggaattcaactga
481 taccaccaccaccatcaccaacaccgccctcagggaaaaaaaaaaggatcaaaaccaga
agtttagaacttctgtgcttagctgaaacgaggggtggacctgggctgggctgggctg
601 ccctcctgtgactgtagcagagcagagattcagtaacaaggtaggggagggtcagggt
attagcaagagaagaaagttaacaaatcctctctcagctctcctgcccagccccaaag
721 cccagaccctgtccactaagcctagctgactctgggaggtgttctgactgaactgaagt
ggccaagaaaggaagtgaagtcagctcagctcagaccctagaaatgaggaaatgttacagaca
841 cactggcaggcaagggaaaccttggccacgtgccacatcagcactcaggaggcagagaca
ggcaggtctctgaaattgaagccagccagtttcttagttccatgccagccatagct
961 acatagtgagaccctgtcccccccccccccaaaaaaaggagctgtgttctctttatca
gtggggccaacagtttaccatgtcccccggaatgaggagatgtgaaggctggcagtgctg

1081 gtgtgtggggcacctgtgcatgaatcctaagctcctctctcaccaccatccagGTGA
V

1141 AGTCCCAGCTTGAATCGAAAGAAAATCAGAAGTTTGATGTCTTTAAAGCCATATCCTTCA
K S Q L E S K E N Q K F D V F K A I S F

1201 AGAGACAGATAGTGGCTGGCACCAACCTCTTCATCAAGgtgggtactgatagtagcttgc
K R Q I V A G T N L F I K

1261 catgaactggggacatagctctcagagtagagcagagtgctcctgcaactcctgcagagaa
ccccttaaggggacatgacatgctctgagaggatgaattgggggtgtaggggtcccgg
1381 ccttaaaggaggagacaaggggttatcactggctaaagttagtgctgggtgctctgtctg
ctcagtttcaagctgggttaagcctggaactggaaccttacctcactcacatgctc

1501 gtctgtctgtctcctccagGTGATGTTGGTGGAGATAAATGCGTGCACCTTGAGGGTGT
V D V G G D K C V H L R V

1561 TTCAACCCCTCCCCATGAAAACAAGCCTTTGACCCTGTCTTCCTATCAGACCAACAAAG
F Q P L P H E N K P L T L S S Y Q T N K

1621 AAAGGCACGATGAGCTCTCTACTTCTGATtcggggctccttgcacctgtgtttgggac
E R H D E L S Y F *

1681 cacgtagtaaaacccccctcggtgatggagagaagggagttgagcagctctctgtgtgat
tctcaactgctgtatcttgcacctccaaataattatcttcagaaaactgtatgatg
1801 atctctctaataatatttttagagactatatgaattgtgctccttagtttaattttt
cccccttggttgaaaaatcacagtcacaccaggcagtggtgggctcacttaatcc

```

**Figure 1** Genomic sequence of the mouse cystatin B gene (GenBank accession no. U59807). Uppercase letters indicate the coding portion of the gene, and lowercase letters designate the 5' flanking region, the 5' UTR, the 2 introns, the 3' UTR, and the 3' flanking region. Amino acids in the cystatin B protein are designated with uppercase letters below the nucleotide sequence. Base pairs are numbered beginning with the ATG start codon. Underlined bases designate the consensus sequence for Sp1-binding sites in the 5' flanking region. Introns 1 and 2 are 1070 and 282 bp in length, respectively. A segment containing six CA repeats in intron 1 is present between nucleotides 1073 and 1087. The stop codon is denoted with an asterisk.

specific to the central nervous system. To understand the relationship between this protein and the symptoms of this disease, we are studying the expression and biochemical properties of the

mouse protein, as well as working toward developing an animal model for EPM1 by disrupting the cystatin B gene in the mouse. As a first step toward these goals, we have isolated and characterized the mouse cystatin B gene. Here we report its structure, mRNA expression patterns, and chromosomal location in the mouse genome.

## RESULTS

## Identification of the Mouse Cystatin B Gene

To identify the mouse cystatin B gene, we designed oligonucleotides based on the known human cystatin B and rat cystatin  $\beta$  cDNA and genomic sequences (Sato et al. 1990, 1992; Pennacchio et al. 1996; K.S. Bhat, unpubl.). One pair of primers (F100/R100; see Methods) yielded a single, ~350-bp long PCR product that is 70% identical to the rat cystatin  $\beta$  genomic nucleotide sequence. We used this primer pair to screen a mouse bacterial artificial chromosome (BAC) genomic library by using the polymerase chain reaction (PCR) and identified BAC clone 310P21 (see Methods). Restriction and hybridization analysis of this BAC clone indicated that the complete mouse cystatin B gene is present in two overlapping DNA fragments: a 6-kb *SacI* fragment and a 1.6-kb *PstI* fragment. We subcloned both of these fragments and determined the nucleotide sequence of the gene and its flanking regions (Fig. 1).

## Nucleotide Sequence Analysis

The genomic sequence of the mouse cystatin B gene is shown in Figure 1. We compared the nucleic acid sequence with that of the human and rat cystatin B ( $\beta$ ) genes and found that all three genes contain 3 exons and 2 introns (Sato et al. 1992; Pennacchio et al. 1996). The mouse open reading frame is contained in all three exons and

## THE MOUSE CYSTATIN B GENE

	1	10	20	30	40	50																																																
mouse	M	M	C	G	A	P	S	A	T	M	P	A	E	T	Q	E	V	A	D	Q	V	K	S	Q	L	E	S	K	E	N	Q	K	F	D	V	F	K	A	I	S	F	K	R	<u>Q</u>	<u>I</u>	<u>V</u>	<u>A</u>	<u>G</u>						
rat	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	I	.	K	.	.	.	.	.	.	.	E	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	V	.	.						
human	.	.	.	.	.	.	.	.	.	.	.	.	Q	.	.	H	I	.	.	.	.	.	.	.	.	R	.	.	E	.	K	.	.	.	.	.	.	.	.	.	.	.	.	V	.	S	.	V	.					
bovine	.	.	.	G	T	.	.	Q	.	.	.	.	.	.	.	A	I	.	.	.	.	.	.	.	.	K	.	.	.	E	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	L	E	.	S	.	L	.		
		60	70	80	90																																																	
mouse	T	N	L	F	I	K	V	D	V	G	G	D	K	C	V	H	L	R	V	F	Q	P	L	P	H	E	N	K	P	L	T	L	S	S	Y	Q	T	N	K	E	R	H	D	E	L	S	Y	F						
rat	.	.	F	.	.	.	.	.	.	E	.	.	.	.	.	E	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	K	.	.	.	T	.			
human	.	.	Y	.	.	.	H	.	.	D	E	D	F	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
bovine	K	.	Y	.	.	.	Q	.	.	D	E	.	D	F	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.

**Figure 2** Alignment of the mouse cystatin B amino acid sequence with those of the rat, human, and bovine sequences. Residues identical to the mouse amino acid sequence are indicated with a dot and residues different from the mouse amino acid sequence are shown with the appropriate amino acid symbol. Boldface amino acid symbols in the rat, human, and bovine sequences represent conservative amino acid changes (positions where different non-polar or charged side chains of similar size appear in the different proteins). The conserved pentapeptide QIVAG at amino acid positions 46 through 50 is underlined.

is 294 bp in length, identical to that found in the human and rat genes. The nucleotide sequence identity between the mouse cystatin B coding sequence and that of the rat and human is 90% and 77%, respectively. The mouse sequence flanking the putative start codon (5'-ACTGCTGCCAAGATG-3') matches the Kozak consensus sequence at most relevant positions (Kozak 1987). The mouse cystatin B gene has an exon-intron structure identical to that of the human and rat genes, and contains GT to AG consensus sequences at its intron/exon junctions (Breathnach and Chambon 1981). The sizes of introns 1 and 2 in the mouse cystatin B gene are 1070 and 282 bp, respectively. In comparison, the sizes of the rat and human cystatin B introns 1 and 2 are 977 and 480 bp, and 1445 and 325 bp, respectively. The putative promoter region of the mouse gene is GC rich and contains no obvious TATAA or CCAAT consensus sequences, similar to promoters of other genes that are expressed ubiquitously (Dyan 1986). This gene has three GC boxes (GGGCGG) identical to those found in other genes where Sp1 transcription factors are known to bind (Pugh and Tjian 1990) (Fig. 1).

### Analysis of the Deduced Mouse Cystatin B Amino Acid Sequence

The mouse cystatin B gene encodes a predicted 98-amino-acid protein, which is identical in length to that of the rat, human, and bovine cystatin B proteins. Alignment of the deduced mouse cystatin B protein sequence reveals 86%, 79%, and 71% amino acid identity and 95%,

88%, and 82% amino acid conservation with the rat, human, and bovine cystatin B protein sequences, respectively (Fig. 2).

Ten of the 11 mammalian stefins whose amino acid sequences were determined before our study contain the sequence QVVAG located approximately in the middle of the primary sequence of the proteins. This pentapeptide comprises the first  $\beta$ -hairpin loop of stefins, and has been shown to interact with the target enzyme in inhibitor/protease complexes (Stubbs et al. 1990b). The first exception to this conservation was bo-

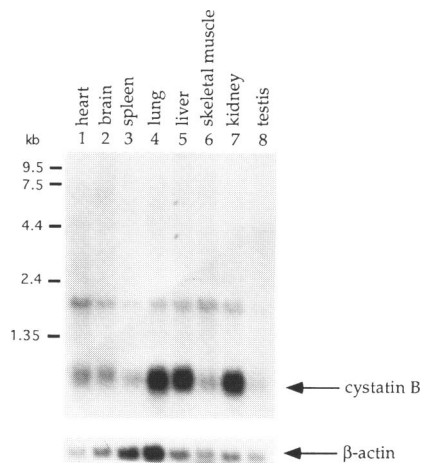
vine cystatin B, which contains the sequence QL-VAG at the equivalent position (Turk et al. 1992). The valine-to-leucine alteration in the bovine protein has little effect on its activity (Turk et al. 1992). Our analysis of the deduced mouse cystatin B amino acid sequence indicates that it contains the segment QIVAG within this region of the protein. Although this change is conservative (valine to isoleucine), mutagenesis studies with the human protein indicate that isoleucine in this position of the pentapeptide results in a fourfold reduction in inhibitory activity against papain (Jerala et al. 1990). Whether this amino acid difference in the mouse protein results in lowered activity remains to be determined.

### Mouse Cystatin B mRNA Expression

To determine the tissue distribution of cystatin B mRNA in the mouse, we hybridized a cDNA probe to a Northern blot containing RNA from multiple tissues (Fig. 3). As an internal standard, we also hybridized a  $\beta$ -actin probe to the same Northern blot. Similar to results seen in humans, all mouse tissues we tested have high levels of cystatin B mRNA at the expected size of ~600 nucleotides. On the basis of normalizing each lane to  $\beta$ -actin signals, heart, liver, and kidney have fivefold higher levels of mRNA than spleen and testis, and twice the levels of mRNA of brain, lung, and skeletal muscle.

### Chromosomal Localization of the Mouse Cystatin B Gene

We determined the chromosomal location of the



**Figure 3** Tissue distribution of mouse cystatin B mRNA. A 218-bp cystatin B cDNA probe was hybridized to an RNA blot purchased from Clontech, Inc. (Palo Alto, CA). Each lane contains 2  $\mu$ g of polyadenylated mRNA from eight mouse tissues. The size of the cystatin B mRNA is  $\sim$ 600 bp, consistent with the mRNA size found in human tissues. In addition, a  $\beta$ -actin probe was hybridized to the same Northern blot to assess the approximate quantity of RNA per lane.

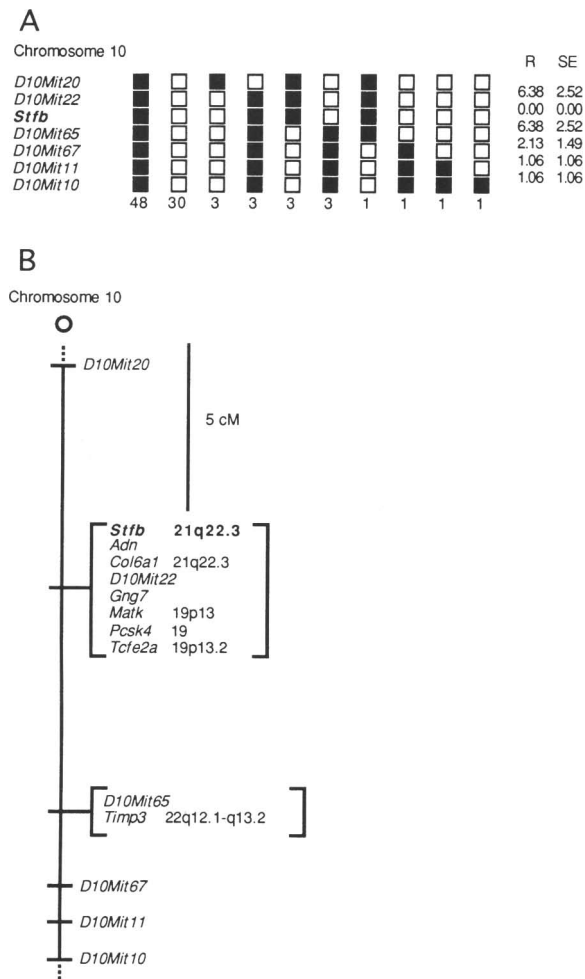
mouse cystatin B gene (*Stfb*) by analyzing the segregation of a polymorphism in the gene in the Jackson BSS panel [(C57BL/6JEi  $\times$  SPRET/Ei) $F_1$   $\times$  SPRET/Ei] interspecific backcross, which has been typed for >2500 loci (Rowe et al. 1994; L. Rowe, pers. comm.). We used oligonucleotides (F100/R100) flanking intron 2 to amplify DNA from the two parental mouse backcross strains *Mus musculus* and *M. spretus*, and detected a *Pst*I restriction fragment length polymorphism in the amplified fragments. We found that *M. musculus* contains a single *Pst*I restriction site in this fragment, whereas *M. spretus* lacks this site. By using this assay to distinguish the two strains, we mapped the mouse cystatin B gene to chromosome 10 (Fig. 4). This pedigree analysis indicated that the mouse cystatin B gene is linked to the loci *D10Mit20*, *D10Mit65*, *D10Mit67*, *D10Mit11*, and *D10Mit10* (Fig. 4A). Furthermore, no recombination events were identified between the gene and the loci *Adn*, *Col6a1*, *D10Mit22*, *Gng7*, *Matk*, *Pcsk4*, and *Tcfe2a* (Fig. 4B). These results establish the order and genetic distances between these loci as *D10Mit20*–(6.38 cM  $\pm$  2.52 s.e.)–[*Stfb*, *Adn*, *Col6a1*, *D10Mit22*, *Gng7*, *Matk*, *Pcsk4*, *Tcfe2a*]–(6.38 cM  $\pm$  2.52 s.e.)–*D10Mit65* (2.13 cM  $\pm$  1.49 s.e.)–*D10Mit67* (1.06 cM  $\pm$  1.06 s.e.)–*D10Mit11* (1.06 cM  $\pm$  1.06 s.e.)–*D10Mit10* (Fig. 4A).

## DISCUSSION

Despite a substantial amount of knowledge about the biochemical properties of the cystatin B protein from several mammalian species, the reasons that defects in the gene lead to EPM1 in humans are not understood. One motivation for isolating and characterizing the mouse cystatin B gene is to provide a means to study the protein and its expression in an experimentally manipulatable mammalian system. In addition, the mouse gene provides a reagent for generating and studying an animal model for the disease. The clones and molecular characterization reported here can be used to construct mouse embryonic stem cells with a completely disrupted cystatin B gene. Because the mutations that we identified in the cystatin B gene in patients with EPM1 may all result in some level of functional protein, it is not yet known whether complete or partial loss of cystatin B activity is the cause of the disease (Pennacchio et al. 1996). It may be necessary to generate mice with these subtle mutations (a splice site mutation and a stop codon mutation at position 67) to produce EPM1-like symptoms, as it is possible that complete loss of function of the gene would be lethal. A successful mouse model of EPM1 will provide a reagent for understanding the progression of symptoms and neuropathology, for developing therapies and for testing hypotheses regarding the roles of particular target proteases in the various symptoms of the disease.

The human cystatin B gene is located on human chromosome 21q22.3, between two syntenic regions of the chromosome that are conserved with portions of two different mouse chromosomes. The synteny in the genomic region proximal to the human cystatin B gene, between  $\alpha$ -A-crystallin (CRYA) and cystathionine- $\beta$ -synthase (CBS), is conserved in the proximal portion of mouse chromosome 17 (Munke et al. 1988; Stubbs et al. 1990a). The synteny in the genomic region distal to the human cystatin B gene, between the liver-type phosphofructokinase (PFKL), the type VI collagen A1 gene (COL6A1), and the  $\beta$  subunit of the Mac-1 leukocyte surface glycoprotein family (S100B), is conserved in a region of mouse chromosome 10 (Justice et al. 1990; MacDonald et al. 1991; Irving et al. 1994). The mapping of the mouse cystatin B gene to mouse chromosome 10 reported here extends the known boundaries of conservation between this region of human chromosome 21q22.3 and the mouse genome.

## THE MOUSE CYSTATIN B GENE



**Figure 4** Mapping of the mouse cystatin B gene to chromosome 10. The gene was mapped to mouse chromosome 10 by interspecific backcross analysis (Rowe et al. 1994). The mouse cystatin B gene locus has been given the symbol *Stfb*, which has been approved by the Mouse Nomenclature Committee. (A) Haplotypes from the Jackson BSS backcross showing a part of chromosome 10 with loci linked to *Stfb*. Loci are listed in order with the most proximal at the top. (■) C57BL6/J/Ei allele; (□) SPRET/Ei allele. The number of animals with each haplotype is given at the bottom of each column of boxes. The percent recombination (R) between adjacent loci is on the right, with the standard error (S.E.) for each R. Missing typings were inferred from surrounding data where assignment was unambiguous. (B) A map of a part of mouse chromosome 10 derived from the Jackson BSS backcross. The map is depicted with the centromere toward the top. Map positions, where known, of the human homologs are listed to the right of the locus names. Raw data from the Jackson Laboratory was obtained from the World Wide Web address <http://www.jax.org/resources/documents/cmdata>. The Mouse Genome Database accession no. for the cystatin B (*Stfb*) gene is MGD-CREX-660. Two groups of loci were not separated by recombination in this study and, therefore, could not be ordered. These are (1) *Stfb*, *Adn*, *Col6a1*, *D10Mit22*, *Gng7*, *Matk*, *Pcsk4*, and *Tcf2a*, and (2) *D10Mit65* and *Timp3*. Mapping was performed with the program Map Manager (Manly 1993).

In addition to providing reagents to produce a mouse model of EPM1, the mouse cystatin B gene can be used to determine whether the gene is involved in aberrant phenotypes in existing mutant mice. The genes for the three mouse mutants *grizzled*, *mocha*, and *jittery* are located on mouse chromosome 10, near the region that contains the cystatin B gene (Taylor et al. 1993). At the current level of resolution of mapping of these genes, it is reasonable to consider the cystatin B gene as a candidate for any of these diseases. In this regard, *jittery* is of particular interest because it displays early onset seizures as part of its phenotype (DeOme 1945), making it a possible mouse model of EPM1. Further mapping of the mouse *jittery* locus relative to the cystatin B gene and mutation analysis, if warranted on the basis of mapping results, will be necessary to test this hypothesis.

## METHODS

### Isolation of Genomic Clones

We isolated BAC 310P21 by PCR screening a pooled mouse genomic BAC library (Research Genetics, Inc., Huntsville, AL; DNA pools were a gift from Greg Lennon, Lawrence Livermore National Laboratory, Livermore, CA) with the oligonucleotides F100 (5'-GGGACAACTACTTCAT-CAAGGT-3') and R100 (5'-GGCTTGTTTCATGAGGGAG-3'), which were designed on the basis of the human cystatin B nucleotide sequence. The PCR conditions were as follows: An initial incubation at 94°C for 5 min followed by 30 cycles for 30 sec at 94°C, 30 sec at 60°C, 1 min at 72°C, with a final incubation for 10 min at 72°C. We subcloned a 6.0-kb *SacI* fragment and a 1.6-kb *PstI* fragment from this BAC clone into plasmid vector pBSISK+ to yield pBS-*SacI*-6 and pBS-*PstI*-1.6, respectively.

### DNA Sequencing

We isolated DNA from overnight cultures by standard al-

## PENNACCHIO AND MYERS

kaline lysis and a Qiagen column. The 6-kb pBS-SacI-6 was digested independently with the blunt end generating restriction enzymes *AluI*, *EcoRV*, *HaeIII*, *HincII*, *HpaI*, *NlaI*, *PvuII*, *RsaI*, and *StuI*. Digests were pooled subsequently and subcloned into the *EcoRV* site of plasmid vector (pBSIISK+) for automated DNA sequencing. We performed sequencing reactions with fluorescently labeled M13-21F and M13R primers and analyzed the reactions with an Applied Biosystems 377A sequencer. We assembled sequence data with the sequence analysis program Sequencher 3.0 (Gene Codes Corporation). PCR products were purified with a Centricon-100 concentrator (Amicon, Beverly, MA) and cloned into a plasmid vector.

## Hybridization Analysis

BAC DNA was digested with restriction enzymes (New England Biolabs) according to the manufacturer's recommendations. Digested DNA was separated by gel electrophoresis and transferred to Hybond-N<sup>+</sup> membranes (Amersham). Membranes were baked for 2 hr at 85°C followed by hybridization with either [ $\alpha$ -<sup>32</sup>P]dATP randomly primed probes or [ $\gamma$ -<sup>32</sup>P]dATP T<sub>4</sub> kinase end-labeled oligonucleotides.

A mouse cystatin B cDNA probe was generated by the random labeling of an RT-PCR product. Primers used for this amplification were F103 (5'-ACGCAGGAGATCGCCGAC-3'), designed from the rat nucleotide sequence, and the R100 primer shown above. Northern blot hybridization was in 5× SSPE, 10× Denhardt's solution, 100 µg/ml salmon sperm DNA, 50% formamide, and 2% SDS at 42°C for 18 hr. Filters were washed with 2× saline sodium citrate (SSC) at room temperature for 30 min and in 0.1× SSC at 60°C for 30 min. All blots were exposed to film overnight with a single intensifying screen at -70°C.

## Mapping of the Mouse Cystatin B Gene

DNA for the BSS interspecific backcross panel was obtained from the Jackson Laboratory. The details of the BSS interspecific backcross are described in Rowe et al. (1994). PCR was performed as described above with primers F100 and R100 on the two parental strains *M. musculus* and *M. spretus*. The 94-animal backcross panel was amplified and subjected to *PstI* restriction digestion. Digested PCR products were separated by 2% agarose gel electrophoresis. *PstI* digestion of the *M. musculus* PCR product resulted in two fragments ~280 and 100 bp in length, whereas the PCR product of the *M. spretus* species was not cleaved by *PstI*, yielding a single fragment ~400 bp in length.

## ACKNOWLEDGMENTS

We thank David Cox, Doug Vollrath, Carlo Iannicola, Christa Prange, Nancy Stone, Virginia Willour, Chris Karlovich, Nila Patil, and members of the Myers laboratory for discussions and support; Lucy Rowe and Mary Barter for assistance in performing the mapping experiments in the mouse backcross panel, Lois Maltais for assistance with nomenclature; and Sami Hussain and Lucia Ramirez for technical support.

The publication costs of this article were defrayed in

part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

## REFERENCES

- Abrahamson, M., A.J. Barrett, G. Salvesen, and A. Grubb. 1986. Isolation of six cysteine proteinase inhibitors from human urine. Their physicochemical and enzyme kinetic properties and concentrations in biological fluids. *J. Biol. Chem.* **61**: 11282-11289.
- Barrett, A.J., H. Fritz, A. Grubb, S. Isemura, M. Jarvinen, N. Katunuma, W. Machleidt, W. Muller-Esterl, M. Sasaki, and V. Turk. 1986a. Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin. *Biochem. J.* **236**: 312.
- Barrett, A.J., N.D. Rawlings, M.E. Davies, W. Machleidt, G. Salvesen, and V. Turk. 1986b. Inhibitors of cysteine proteases. In *Proteinase inhibitors* (ed. A.J. Barrett and G. Salvesen), pp. 515-569. Elsevier, Amsterdam, The Netherlands.
- Berkovic, S.F., F. Andermann, S. Carpenter, and L.S. Wolfe. 1985. Progressive myoclonus epilepsies: Specific causes and diagnosis. *N. Eng. J. Med.* **315**: 296-305.
- Breathnach, R. and P. Chambon. 1981. Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50**: 349-383.
- DeOme, K.B. 1945. A new recessive lethal mutation in mice. *Univ. Calif. Publ. Zool.* **53**: 41-66.
- Dynan, W.S. 1986. Promoters for housekeeping genes. *Trends Genet.* **2**: 196-197.
- Irving, N.G., D.E. Cabin, D.A. Swanson, and R.H. Reeves. 1994. Gene order is conserved within the human chromosome 21 linkage group on mouse chromosome 10. *Genomics* **21**: 144-149.
- Jerala, R., M. Trstenjak-Prebenda, L. Kroon-Zitko, B. Lenarcic, and V. Turk. 1990. Mutations in the QVVAG region of the cysteine proteinase inhibitor stefin B. *Biol. Chem. Hoppe-Seyler* **371**: 157-160.
- Justice, M.J., L.D. Siracusam, D.J. Gilbert, N. Heisterkamp, J. Groffen, K. Chada, C.M. Silan, N.G. Copeland, and N.A. Jenkins. 1990. A genetic linkage map of mouse chromosome 10: Localization of eighteen molecular markers using a single interspecific backcross. *Genetics* **125**: 855-866.
- Kozak, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15**: 8125-8148.
- MacDonald, G., M.L. Chu, and D.R. Cox. 1991. Fine structure physical mapping of the region of mouse chromosome 10 homologous to human chromosome 21. *Genomics* **11**: 317-323.

## THE MOUSE CYSTATIN B GENE

- Manly, K.F. 1993. A Macintosh program for storage and analysis of experimental genetic mapping data. *Mamm. Genome* **4**: 303–313.
- Munke, M., J.P. Kraus, T. Ohura, and U. Francke. 1988. The gene for cystathionine beta-synthase (CBS) maps to the subtelomeric region on human chromosome 21q and to proximal mouse chromosome 17. *Am. J. Hum. Genet.* **42**: 550–559.
- Pennacchio, L.A., A.-E. Lehesjoki, N.E. Stone, V.L. Willour, K. Virtaneva, J. Miao, E. D'Amato, L. Ramirez, M. Faham, M. Koskiniemi, J.A. Warrington, R. Norio, A. de la Chapelle, D.R. Cox, and R.M. Myers. 1996. Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1). *Science* **271**: 1731–1734.
- Pugh, B.F. and R. Tjian. 1990. Mechanism of transcriptional activation by Sp1: Evidence for coactivators. *Cell* **61**: 1187–1197.
- Rawlings, N.D. and A.J. Barrett. 1990. Evolution of proteins of the cystatin superfamily. *J. Mol. Evol.* **30**: 60–71.
- Ritonja, A., W. Machleidt, and A.J. Barrett. 1985. Amino acid sequence of the intracellular cysteine proteinase inhibitor cystatin B from human liver. *Biochem. Biophys. Res. Comm.* **131**: 1187–1192.
- Rowe, L.B., J.H. Nadeau, R. Turner, W.N. Frankel, V.A. Letts, J.T. Eppig, M. Ko, S.J. Thurston, and E.H. Birkenmeier. 1994. Maps from two interspecific backcross DNA panels available as a community genetic mapping resource. *Mamm. Genome* **5**: 253–274.
- Sato, N., K. Ishidoh, Y. Uchiyama, and E. Kominami. 1990. Molecular cloning and sequencing of cDNA for rat cystatin beta. *Nucleic Acids Res.* **18**: 6698.
- . 1992. Structural organization of the gene encoding rat cystatin beta. *Gene* **114**: 257–260.
- Stubbs, L., J. Kraus, and H. Lehrach. 1990a. The alpha-A-crystallin and cystathionine beta-synthase genes are physically very closely linked in proximal mouse chromosome 17. *Genomics* **7**: 284–288.
- Stubbs, M.T., B. Laber, W. Bode, R. Huber, R. Jerala, B. Lenarcic, and V. Turk. 1990b. The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *EMBO J.* **9**: 1939–1947.
- Taylor, B.A., W.N. Frankel, M. Burmeister, and E. Bryda. 1993. Encyclopedia of the mouse genome III. October 1993. Mouse chromosome 10. *Mamm. Genome* **4**: S154–S163.
- Turk, B., I. Krizaj, and V. Turk. 1992. Isolation and characterization of bovine stefin B. *Biol. Chem. Hoppe-Seyler* **373**: 441–446.
- Turk, V. and W. Bode. 1991. The cystatins: Protein inhibitors of cysteine proteinases. *FEBS Letter* **285**: 213–219.

Received July 23, 1996; accepted in revised form August 22, 1996.